An experiment tested the efficacy of "discovery learning" in a school setting. Fifty-two bright children ages one to seven were randomly divided into two groups: a discovery learning, non-directed group and a guided learning, teacher-directed group. After both groups had spent eight months using an individualized "talking typewriter" program designed to teach spelling and reading skills, the children were tested on the WRAT (Wide Range Achievement Test) reading and spelling achievement tests. The two groups showed no significant differences. Objective and subjective measures of motivation likewise yielded no strong difference between the two groups. It thus appears that any differences between the two teaching approaches either do not exist or are so small as to be mitigated in long-term school use. (RB)
Final Report

Project No. 1F041
Contract No. OEC-6-71-0484-(509)

Donald S. Baer
Creative Learning Center
1616 E. Illinois
Dallas, Texas 75216

COMPARING TWO METHODS OF TEACHING PRE-SCHOOL CHILDREN TO READ AND SPELL AT AN ELECTRIC TYPEWRITER, THEIR REACTION TO THE EXPERIENCE, AND THE SIGNIFICANCE TO EARLY CHILDHOOD EDUCATION

April 1972

U.S. DEPARTMENT OF HEALTH, EDUCATION, AND WELFARE
Office of Education
Cooperative Research Program
(Regional Project Research)
Final Report

Project No. 1F041
Contract No. OE5-6-71-0414-(509)

Donald S. Baer
Creative Learning Center
1616 E. Illinois
Dallas, Texas 75216

COMPARING TWO METHODS OF TEACHING PRE-SCHOOL CHILDREN TO READ AND SPELL AT AN ELECTRIC TYPEWRITER, THEIR REACTION TO THE EXPERIENCE, AND THE SIGNIFICANCE TO EARLY CHILDHOOD EDUCATION

April 1972

U.S. DEPARTMENT OF HEALTH, EDUCATION, AND WELFARE
Office of Education
Cooperative Research Program
(Regional Project Research)
SUMMARY

By testing the efficacy of "discovery learning" in a school setting it may be possible to avoid some of the irrelevancy, lack of comparability, and other such pitfalls present in many of the studies in the literature. In addition, such an in-school experiment allows investigation of other aspects of discovery learning versus other methods of instruction -- especially the effects of the various methods on the motivation of the children.

Fifty-two bright children ages one to seven, of various racial and social-economic backgrounds, were pre-tested on the WRAT reading and spelling achievement subtests. They were then randomly divided into two groups: a discovery learning, non-directed group and a guided learning teacher-directed group. After eight months of an individualized "talking typewriter" program designed to help the children acquire spelling and reading skills, the children were post-tested on the WRAT and questioned to determine their subjective reactions to the program.

When compared for gains in achievement the two groups showed no significant differences. Objective and subjective measures of motivation likewise yielded no strong differences between the two groups. It is suggested that advantages claimed for one or the other teaching methods are mitigated in long-term school use.
Final Report

Project No. 1F041
Contract No. OEC-6-71-0484-(509)

COMPARING TWO METHODS OF TEACHING PRE-SCHOOL CHILDREN TO READ AND SPELL AT AN ELECTRIC TYPEWRITER, THEIR REACTION TO THE EXPERIENCE, AND THE SIGNIFICANCE TO EARLY CHILDHOOD EDUCATION

Donald S. Baer

Creative Learning Center

Dallas, Texas

March 1972

The research reported herein was performed pursuant to a contract with the Office of Education, U.S. Department of Health, Education, and Welfare. Contractors undertaking such projects under Government sponsorship are encouraged to express freely their professional judgment in the conduct of the project. Points of view or opinions stated do not, therefore, necessarily represent official Office of Education position or policy.

U.S. DEPARTMENT OF HEALTH, EDUCATION, AND WELFARE

Office of Education
Cooperative Research Program
(Regional Project Research)
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td>1</td>
</tr>
<tr>
<td>Methodology</td>
<td>5</td>
</tr>
<tr>
<td>Results</td>
<td>9</td>
</tr>
<tr>
<td>Conclusions</td>
<td>15</td>
</tr>
<tr>
<td>Bibliography</td>
<td>16</td>
</tr>
<tr>
<td>Appendices</td>
<td>18</td>
</tr>
</tbody>
</table>
TABLES

Table 1: Chi-square Data Relating to Hypothesis 2A 12
Table 2: T-test Data Relating to Hypothesis 2B 12
Table 3: T-test Data Relating to Hypothesis 2C 13
Table 4: Chi-square Data Relating to Hypothesis 2D 13
Table 5: Chi-square Data Relating to Hypothesis 2E 14

FIGURES

Figure 1: T-test Differences in Semantic Differential Ratings for the ND Versus TD Groups 10
INTRODUCTION

Several factors account for the recently renewed interest in the "discovery learning" variable. The public schools have come under increased scrutiny by many social groups, and changes in curriculum have resulted. There has been a spread of the "open classroom" and growing interest in the "free school" movement. An increased interest in Montessori principles -- especially regarding her emphasis on greater freedom on the part of the child -- has come about. Montessori was convinced that a teacher must "never substitute his own intelligence for that of the child, but rather make the child himself think, and induce him to exercise his own activity." (8:44) In addition, educational theorists -- especially Bruner -- have brought the issue to public attention again, after a lapse of many years.

Early educational theoreticians -- Rousseau, Montessori, and Dewey -- emphasized the efficacy of discovery learning as a method preferable to the highly structured urban classroom settings. In fact, several early articles in this field tended to contrast "discovery" learning with "rote" learning; as has been noted by Wittrock (15:44) few people would be likely to vote against the superiority of the former method, given those labels.

Unfortunately, for many years the literature on discovery learning has been crowded with biased labels and ostensibly plausible hypotheses based on anecdotal, analogic, or hypothetical concepts. Many of them apparently make good sense, but lack easy testability. An example, from a non-educational context:

For reasons never clearly specified, it is argued by some ... that whatever is discovered (by a person) is intrinsically more valuable than whatever (is learned) due to ... attempts to teach him .... An analogy of approach is that of a man who pushes his stalled car to a gas station and is greeted by a mechanic who, believing in self-discovery, insists
that his client deduce the cause and remedy for his car's difficulty. Under the guidance of a good mechanic, the man will probably arrive at a correct diagnosis, and even be able to repair the car, but the joy of self-discovery will hardly compensate him for the expenditure in time and fees that this approach implies. (13:102)

Similar analogies, hyperbole, etc. are used also to bolster the case for discovery learning, both in and out of the education context. Clearly, what is needed in such areas of dispute is research to test the principles involved. There is hope that the concepts can be accurately defined, that specific independent and dependent variables can be formulated and results compared, and that the results can be stated precisely and be checked with appropriate statistical procedures. In that way, the argument might be satisfactorily decided on the basis of fact.

Indeed, many studies have been undertaken in an attempt to isolate the "discovery" mode of learning, so that comparisons between it and other methods of learning might be made.* Hermann (3) has tried to summarize the results; it is a noble attempt, but the data being compared are often based on a wide range of concepts that were tested under varying experimental conditions with subjects varying widely in ages, abilities, etc. As a result, even a careful reading of the Hermann article leaves one puzzled as to whether the benefits or disadvantages of the various learning methods have actually been shown with any certainty. The article is, in fact, a

*Together, the Hermann (3) and Wittrock (15) articles summarize the majority of the work that has been done on the "discovery learning" method. Hermann's paper pulls together the major experimental findings, and investigates several subsidiary variables. Wittrock's paper analyzes the experiments, explains underlying issues, and discusses various research approaches. The book in which the Wittrock article appears -- Shulman and Keislar (12) -- has a comprehensive bibliography of the discovery learning literature and also contains other pertinent articles.
valiant effort to make sense out of a mass of data that are not really unified, but are a series of varied approaches to different subparts of the whole problem.

Wittrock, the man who has done the most work in attempting to synthesize the data -- Kersh and Wittrock (6), Wittrock (14), Wittrock (15) -- has pointed out that there are simply too many unsolved problems in this area of research. He notes the large variety of dependent variables used, the basic flaws in the research techniques, and comments:

The reader is forewarned that the current state of research on discovery is very disappointing and precludes any important conclusions about teaching or learning.... The literature on learning by discovery does not lend itself to terse summaries. The studies are seldom closely interrelated to one another. The procedures are sometimes naive and crude and evidence only the researcher's preference for a type of treatment. Only by summarizing a few studies in detail can one appreciate the several meanings of the learning by discovery hypothesis, the severe shortcomings of the research, the futility of an attempt to gloss over the particulars of a study, and the meaninglessness of generalizations based on these studies. (15:45)

There is, however, another route to the collection of data about discovery learning. Bruner has stated at the conclusion of his outline of the discovery learning hypothesis that the hypothesis is "of such important human implications that we cannot afford not to test it -- and testing will have to be in the schools." (1:26) This opinion is echoed by Keislar and Shulman: "Generalizing from a two-month experiment to a two year curriculum may involve overextrapolation, but at
least it seems more reasonable than to base one's judgment on data collected from a fifty minute laboratory session on the same problem." (4:196)

Basically, these men are suggesting that a naturalistic -- in school -- study may be more likely to produce answers to problems about how children learn, than will a more narrowly-defined but less relevant experimental study. In addition to being able to investigate the discovery learning hypothesis in a natural setting, a study done in a school can also investigate several other related questions that cannot be investigated in the laboratory. Specifically: Friedlander (2) believes that children given the correct answer will probably tend, after a number of such experiences, to have self-motivation eliminated. Similarly, Kersh (5) claimed to have found, from observational reports, that the "no help" group of children were more motivated to learn on their own; he suggested that a choice existed between getting maximum understanding or getting maximum motivation. Furthermore, Kornreich (7) has pointed out the need to study differing teaching methods to determine: specification of just what the differences in methods really amount to, analysis of changes in student behavior as a function of the differences in teaching methods, and differences in criterion performances (learning) among students in the various teaching situations.

In sum, there clearly exists a need for "in school" investigations of the discovery learning hypothesis for at least two reasons. First of all, such investigations can provide data to support or reject the use of discovery learning methods; minimally, such research projects are likely to provide data that will suggest new areas of the problem that ought to be considered further. Secondly, such investigations can provide information relevant to several side issues -- specifically those hypotheses that have been set forth concerning the likely side effects of discovery learning versus other types of learning.
Setting and subjects:
The Creative Learning Center in Dallas, Texas is a year-round Montessori day school for bright children aged one to seven; seventy percent of the students are from poverty-area homes, and the children come from Black, White, Mexican-American, and American Indian families. Students are pre-selected for admission to the school by various social agencies, are then tested on standard tests, and are admitted if they are found to be above average in intelligence. During the course of the study, fifty-five children were registered as students at the school. Two children did not attend long enough to be given complete testing, and one child would not come to the research room. The data from the remaining fifty-two children were used in the study.

The learning situation:
It was necessary to set up a learning situation where the effects of discovery learning could be compared to the effects of another teaching method. Given the ages of the children involved, and the availability of standardized tests, it was decided to teach reading and spelling skills using a "talking typewriter" program similar to that originally designed by Moore (9), and incorporating some of the non-computerized variations suggested by Nimnicht (10). Basically, the system consisted of one child and a teacher working together, with the child doing the typing and the teacher aiding the child in various ways. The teacher and child were seated next to each other, the child in front of an electric typewriter modified for school use by the addition of lower-case letters on the keyboard and the choice of a very large, easy-to-read type size. Other materials included masonite letters, cards with pictured objects and the object names printed beneath, a typed list of the names of the children in the school, some early reading books, and a typed alphabet. Additional supplies were brought in as dictated by the progress of the children.
Pre-typed lesson plans were created that called for a wide range of typing activities such as letter and word copying, word completion, sentence completion, etc.

Besides allowing great control over the teaching method due to only one child being present at a time, this system made videotaping for research purposes easy. Also, with the use of a carbon paper system, copies of all the typing done by the children could be kept, available for later analyses. The researcher was new to the field of education and had no particular preference for any one teaching method; this was especially important in that the researcher also filled the role of teacher in the study.

Independent variable:

The children were randomly divided into a non-directed (ND) group and a teacher-directed (TD) group; each group contained 26 children. Except for the amount of teacher direction given, all other aspects of the learning situation were kept as similar as possible for the two groups. Specifically, the children were asked about coming to the typing room in a random order, and all were approached with the same question: "Would you like to come to the typing room?" (If they refused, they were asked again the next day. If they refused again, they missed that turn. The children averaged slightly more than one turn per week.) All were given the same maximum amount of time in the room (10 minutes), were allowed to leave whenever they desired, and were given a choice of typing on a pre-planned lesson or on a blank piece of paper. No matter which paper was chosen, the teacher inserted it in the machine, due to the difficulty the children had manipulating two sheets of paper and the carbon. For both groups, if a child hit more than one key at a time the machine was shut off for ten seconds, the jammed keys were released, and the teacher said, "Just hit one key at a time." Throughout the typing session, the teacher tried to maintain good rapport with the child, and at the end of the time the teacher quickly summarized what had been done during that lesson.
The two groups differed in the following ways: For the NDs, no explanation was given about the materials on the pre-planned lesson sheet unless the children indicated, verbally or otherwise, that they wanted an explanation. The teacher did not urge NDs to do any particular kind of work, nor did the teacher mention or attempt to correct any mistakes made by the ND children unless the children asked for the information. No attempt was made to keep the NDs at the typewriter; several of them spent time looking at pictures, playing with the videotape camera, etc. Basically, every effort was made to put the NDs in a discovery learning situation, where the initiation and direction of learning tasks were left up to the children.

Children in the TD group were told the purpose of the pre-planned lesson sheets (for example, "... and on this paper, the idea is for you to make another letter just like the one you see here.") before they chose between it and the blank sheet of paper. If TDs chose the pre-planned lesson sheet, it would be put in the typewriter, and they would then be reminded of the purpose of the particular lesson and would be urged gently to work at it in that way. If TDs chose to work independently or to look at other things in the room, they were encouraged at least three times to work at the lesson. Strong pressure by the teacher was avoided. For this group, errors in copying or writing words and letters were mentioned, and the children in this group were encouraged to correct their mistakes. For the TDs, the goal was to set up a learning situation in which the teacher lead each child through progressive stages of reading and spelling achievement.

Test instruments:

The Wide Range Achievement Test was given to the children at the beginning and end of the study, the average time between pre-testing and post-testing being 7 months and 26 days. All tests were scored double-blind. The major dependent variables, hereafter called "gain scores," equaled the differences between the
child's post-test and pre-test scores in reading and spelling. At the end of the study, the children were given an oral questionnaire (Appendix A) in order to ascertain their subjective opinions of the typing program.

Hypotheses:

1. Gain scores for the NDs will differ significantly from gain scores for the TDs -- a two-tailed hypothesis. The size of the difference may relate to the children's ages, sexes, social-economic statuses (poverty versus non-poverty), or races.

To test the Friedlander contention that children in the TD mode will have lessened self-motivation and the Kersh contention that NDs will learn more on their own, and to explore the ramifications of Kornreich's suggestion that children in different teaching systems will behave differently, these one-tailed hypotheses were formulated:

2A. The TDs will be less willing than the NDs to come to the typing room;
2B. The TDs will do less work than the NDs;
2C. The TDs will choose to do pre-planned lessons (versus blank paper) a smaller percentage of the time than will the NDs;
2D. The TDs will leave the typing room early more often than will the NDs; and
2E. The TDs will indicate on their answers to the questionnaire less pleasure in typing, less desire to come to the typing room, and less desire to work at the typewriter (versus working on other things in the typing room) than will the NDs.

A supplementary evaluation of the value of the "talking typewriter" program can be made by checking Nimnicht's contention that there was "a positive correlation of .76 between the number of times a child went to the booth and the (learning) phase the child reached." (10:37) The third hypothesis is, therefore, that:

3. There will be a positive correlation between the number of lessons given and the size of the gain scores.
RESULTS

Validity of independent variable:

In attempting to prove the validity of the experimental manipulation -- that is, to prove that the ND and TD groups actually were taught under differing conditions -- a videotape was made of the opening minute of typing sessions with each of 42 randomly chosen children. This tape, rating instructions, and semantic differential rating sheets (Appendix B) were given to each of three raters -- all familiar with the field of education and professionally involved in it. The raters were not told the purpose of the study, nor were they told which children were in each of the groups.

The fifteen pairs of adjectives on the semantic differential were chosen so that three pairs reflected the manipulations present in the independent variable (teacher-led:child-led, directive:non-directive, and leading:following). In choosing other pairs of adjectives, emphasis was placed on finding those pairs that might be likely to eliminate alternate hypotheses if learning differences were found between the NDs and TDs. For example, it could easily be argued that different levels of attention by the teacher might be present in teaching one group with a directive method, and the other group with a non-directive method; therefore, one pair of adjectives included was "attentive:negligent." For the most part, the remaining pairs of adjectives were chosen from the major factors found by Osgood, Suci, and Tannenbaum (11:53-61).

Of the fifteen scales, it was predicted that only the three scales reflecting the independent variable manipulation would show significant differences between the two groups of children. It was expected that: the teacher would be rated as more "leading" for the TD group; the teacher would be rated as more "directive" for the TD group; the lesson would be rated as more "teacher-led" for the TD group. All other scales were predicted not to differ significantly for the two groups. Reliability of the semantic differential has been previously demonstrated (11).

Rating results may be seen in Figure 1. The
Figure 1
T-test Differences in Semantic Differential

predicted-difference scales
Ratings for the ND Versus TD Groups

no-predicted-difference scales

Rater 1
Rater 2
Rater 3

ive: attentive: orderly: helping:
active negligent disorderly hindering
above child's level: neat: distressed: pleasant:
below child's level: messy comfortable unpleasant

Rating Scales
three raters agreed -- all at the p<.01 level for two of the three predicted-difference scales, and with one rater at the p<.01 level and two raters at the p<.10 level for the third scale -- that the two groups of children differed in the expected directions. On only one of the twelve no-predicted-difference scales did as many as two of the raters agree that there were significant differences between the two groups of children; on 6 no-predicted-difference scales only one rater found significant differences; on 5 no-predicted-difference scales, no rater found significant differences. The only no-predicted-difference scale on which as many as two raters found significant differences (one at the p<.05 and one at the p<.10 level) was the "active:passive" scale. Both raters rated the teacher as more active with the TD group. This finding offers little reproach to the method, simply indicating that directiveness on the part of a teacher requires more activity than does non-directiveness.

The results would be "cleaner" if all of the no-predicted-difference scales had shown no significant differences for any rater, but in light of the few raters used and the relatively broad ratings that were called for, it is clear that the experimental manipulations have been shown to be valid. It is possible for a single teacher to structure two methods of teaching -- one directive and the other non-directive -- so that valid and predictable differences can be found between them; furthermore, this can be done without producing strong differences in areas other than the one being manipulated.

Data for the hypotheses:
The validity of the experimental manipulation having been proven, consider the data for the major hypotheses of the study. For the first hypothesis, no significant differences in gain score existed between the TDs and NDs, either when the scores were compared alone, or when they were controlled for age, sex, social-economic status, or race of the children.

For hypothesis 2A, that the TDs will be less willing than the NDs to come to the typing room; the data in Table 1 indicate that no such
difference existed at a significant level.

Table 1
Chi-square Data Relating to Hypothesis 2A

<table>
<thead>
<tr>
<th>No refusals to come typing</th>
<th>TD</th>
<th>ND</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>17</td>
<td>13</td>
</tr>
<tr>
<td>At least one refusal to come typing</td>
<td>9</td>
<td>13</td>
</tr>
</tbody>
</table>

\[ \chi^2 = 1.26 \text{ (one-tailed)} \]

\[ p > .10, \text{ non-significant} \]

For hypothesis 2B, that the TDs will do less work than will the NDs, it must be noted that the amount of work done by the children could be measured in many ways. Due to the time-consuming nature of attempting to count "number of words typed" or "number of letters typed" and due to the questionable validity of such measures -- because many children practiced typing numbers, practiced reading, and practiced various mechanical operations on the typewriter -- it was decided that a comparison of amount of work done between the TDs and NDs would be based on the number of pages typed per lesson. Table 2 shows the pertinent data.

Table 2
T-test Data Relating to Hypothesis 2B

Average number of pages per lesson:

<table>
<thead>
<tr>
<th></th>
<th>TD</th>
<th>ND</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1.14</td>
<td>1.25</td>
</tr>
</tbody>
</table>

\[ t = 1.59 \text{ (one-tailed)} \]

\[ p < .10, \text{ trend} \]
For hypothesis 2C, that the TDs will choose to do pre-planned lessons (versus blank paper) a smaller percentage of the time than will the NDs; Table 3 shows that no such difference existed at a significant level.

Table 3
T-test Data Relating to Hypothesis 2C

<table>
<thead>
<tr>
<th>Percentage of lessons chosen, (versus blank paper)</th>
</tr>
</thead>
<tbody>
<tr>
<td>TD</td>
</tr>
<tr>
<td>.27</td>
</tr>
<tr>
<td>( t = 1.18 ) (one-tailed)</td>
</tr>
<tr>
<td>( p &gt; .10 ), non-significant</td>
</tr>
</tbody>
</table>

For hypothesis 2D, that the TDs will leave the typing room early more often than will the NDs, Table 4 shows that the difference between the groups was non-significant.

Table 4
Chi-square Data Relating to Hypothesis 2D

<table>
<thead>
<tr>
<th>Child never left room before time was up</th>
<th>TD</th>
<th>ND</th>
</tr>
</thead>
<tbody>
<tr>
<td>Child left at least once before time was up</td>
<td>13</td>
<td>18</td>
</tr>
</tbody>
</table>

\( \chi^2 = 2.00 \) (one-tailed)
\( p > .10 \), non-significant

For hypothesis 2E, the children's answers to seven of the eight questionnaire items showed acquiescence or reticence to great to make the items reliable. The answers to question 3 -- "What do you like to do best in the typing room?" -- were divided so that all replies indicating
preference for spelling, reading, or typing activities were included under the "typing" category, and all other responses were counted in the "non-typing" category. These categories were decided upon before detailed data analyses were done; note also that because some of the subjects gave more than one preference, the total number of preferences is greater than the number of children in the study. Table 5 supplies the relevant data.

Table 5
Chi-square Data Relating to Hypothesis 2E

<table>
<thead>
<tr>
<th></th>
<th>TD</th>
<th>ND</th>
</tr>
</thead>
<tbody>
<tr>
<td>Typing</td>
<td>14</td>
<td>19</td>
</tr>
<tr>
<td>Non-typing</td>
<td>12</td>
<td>11</td>
</tr>
</tbody>
</table>

$X^2 = .56$ (one-tailed)
p > .10, non-significant

In checking Nimnicht's finding of a .76 correlation between the number of visits to the typing room and the learning phase reached, a Pearson product-moment correlation was run between the number of visits made by the children and the size of the combined -- reading and spelling -- gain scores. An r of +.40 resulted, p < .01; this strongly supports Nimnicht's findings.
CONCLUSIONS

Despite ample evidence that the TD and ND groups of children were taught under two distinctly different teaching methods, differences between the groups in reading and spelling achievement were wholly non-significant. While it is unfortunate to find support only for the null hypothesis, such a finding suggests that the advantages often claimed for one or the other method are rather overstated. Even if it is the case in short-term research studies that TD groups do differ from ND groups in amount of achievement, the results from the present study indicate that such differences are unlikely to carry over to long-term, school-like situations.

As far as subsidiary hypotheses are concerned; while a single measure of motivation did show the TDs to differ from the NDs at the p<.10 level, this trend is almost completely lost among the four other measures of motivation which yielded no significant differences between the two groups. Clearly, the thrust of the data does not support either Friedlander or Kersh in their expectations of motivational differences between the teacher-directed and the non-directed groups of children.

Finally, as a subsidiary point, further evidence of the efficacy of a "talking typewriter" program with young children was found.
BIBLIOGRAPHY


Appendix A

Post-test questionnaire (administered orally)

1. Do you like to come to the typing room?

2. If you could change anything about the way we play with the typewriter, what would you change?

3. What do you like to do best in the typing room?

4. Do you have a typewriter at home? Do you ever use it? How often? Who else uses it?

5. Would you like to come to the typing room more often?

6. What do you not like about the typewriter?

7. Do you like to type?

8. What do you like best about the typewriter?
Appendix B

Semantic differential rating sheet (form 1)*

Typewriter Project, Page __

Subject Number __

A. The instructor was:

- FOLLOWING / / / / / / / LEADING
- DIRECTIVE / / / / / / / NON-DIRECTIVE
- COMFORTABLE / / / / / / / DISTRESSED
- HINDERING / / / / / / / HELPING
- DISAPPROVING / / / / / / / PRAISING
- ATTENTIVE / / / / / / / NEGLECTED
- RESPONSIVE / / / / / / / NON-RESPONSIVE
- ACTIVE / / / / / / / PASSIVE
- UNPLEASANT / / / / / / / PLEASANT

B. The materials were:

- NEAT / / / / / / / MESSY
- ABUNDANT / / / / / / / SCARCE
- APPROPRIATE / / / / / / / INAPPROPRIATE

C. The lesson was:

- ABOVE
- CHILD'S LEVEL / / / / / / / CHILD'S LEVEL
- DISORDERLY / / / / / / / ORDERLY
- TEACHER-LED / / / / / / / CHILD-LED

BELOW

*Form 2, the same as form 1 but with the scales in another order, was alternated with form 1 in the stack of 42 rating sheets given to each rater. This was done to lessen any error due to scale position.